Available online at www.sciencedirect.com

ScienceDirect

Journal of Multivariate Analysis 98 (2007) 1356–1375

Journal of Multivariate Analysis

www.elsevier.com/locate/jmva

Moving estimates test with time varying bandwidth

Okyoung Na, Sangyeol Lee*

Department of Statistics, Seoul National University, Republic of Korea

Received 29 September 2005
Available online 28 February 2007

Abstract

In this paper, we consider the problem of testing for parameter changes in time series models based on a moving estimates (ME) test. It is widely accepted that detecting some changes, for instance, those caused by temporary parameter shifts by the existing cusum test is difficult. A MV test with a fixed bandwidth has been developed to circumvent the defect, but the test still does not perform well under certain conditions. Motivated by this, we propose a MV test with a time varying bandwidth to outperform the original test. In order to illustrate our findings, we have provided simulation results.

AMS 1991 subject classification: 62G10; 62G20; 62M10

Keywords: Test for parameter change; Moving estimates test; Cusum test; Time series; Strong mixing processes

1. Introduction

The problem of testing for parameter changes in statistical models has a long history. It was originally studied in the quality control context and was subsequently extended to various areas such as economics, finance, medicine, and seismic signal analysis. For a general review of the change point problem, see the articles appearing in [5]. The cusum test has long been popular for testing for the existence of change points and then allocating them. See, for instance, Picard [18], Inclán and Tiao [8], Jandhyala and MacNeill [9], and Tang and MacNeill [16]. An advantage of using the cusum test is that there are no assumptions imposed on the underlying distribution of observations unlike in the other methods, for instance, the parametric approach which is not suitable to the test a change in the autocorrelations of stationary time series. Lee et al. [12], Lee and Na [14], and Lee et al. [13] used the cusum test to overcome this problem and applied it

* Corresponding author.
E-mail addresses: okyoung0@hanmail.net (O. Na), sylee@stats.snu.ac.kr (S. Lee).

0047-259X/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2007.02.003

to various cases such as the test for parameter changes in random coefficient autoregressive and autoregressive conditional heteroscedastic (ARCH) models.

Although the cusum test is very useful in actual practice, it has a drawback in that it loses efficiency when the change point is not located in the middle of the time series. Further, it does not effectively detect certain type of changes such as the temporary parameter shift as seen in [4]. To remedy this, they proposed a class of moving estimates (ME) tests. Their simulation result demonstrates that in the case of double changes, the ME test with moving window bandwidth h $(0 < h < 1)$ is superior to or comparable to other competing tests such as the AVG-F and EXP-F tests. However, when the interval between shifts is short or a periodic parameter change occurs, the ME test has a tendency to produce low powers. Therefore, in this study, we consider a new ME test with the moving window bandwidth $h = h_n$, where h_n is a sequence of positive real numbers decaying to 0 as $n \to \infty$.

This paper is organized as follows. In Section 2, we introduce the MV test with a time varying bandwidth for a mean change in an i.i.d. sample. In Section 3, we extend the result in Section 2 to a time series case, and discuss a method for allocating the change points. In Section 4, we report the simulation results in order to evaluate the performance of our tests and demonstrate that our tests perform adequately. In Section 5, we provide the proofs of the theorems presented in Section 3.

2. ME test for a mean change in i.i.d. sample

In this section we introduce the ME test with a time varying bandwidth to test for the constancy of the means in i.i.d. observations. Chu et al. [4] considered the model

$$y_i = \mu_i + \epsilon_i, \quad i = 1, \ldots, n,$$

where $\epsilon_1, \ldots, \epsilon_n$ are i.i.d. r.v.'s with mean zero and variance one, and proposed the ME test for testing

H_0: $\mu_i = \mu_0$ for all $i = 1, 2, \ldots, n$, where μ_0 is assumed to be unknown vs.
H_1: μ_i varies over time i

based on the estimates of μ_0

$$\hat{\mu}_k = \frac{1}{[nh]} \sum_{i=k+1}^{k+[nh]} y_i, \quad k = 0, 1, \ldots, n - [nh],$$

with a fixed bandwidth h, where $[nh]$ denotes the integral part of nh. This test has an advantage over the ordinary cusum test in that the former outperforms the latter in some situations (see [4,10], and the simulation results in Section 4). However, there is a possibility that the fixed h may not be sufficiently good for the test to perform efficiently and much smaller h is preferred as it yields more accurate results. Following this reasoning, we consider the ME test where h varies with the sample size n. More precisely, $h := h_n$ is a sequence of positive real numbers satisfying

$$h \to 0 \quad \text{and} \quad nh \to \infty \ \text{as}\ n \to \infty. \tag{1}$$

In essence, the ME test functions by detecting significant changes in the fluctuations of the MV $\hat{\mu}_k - \hat{\mu}$, where $\hat{\mu} = n^{-1} \sum_{i=1}^{n} y_i$, which naturally leads us to employ the maximum type test statistic

$$\max_{0 \leqslant k \leqslant n-[nh]} \left| \hat{\mu}_k - \hat{\mu} \right|, \tag{2}$$

or its quadratic version

$$\sum_{k=0}^{n-[nh]} \left(\hat{\mu}_k - \hat{\mu}\right)^2. \tag{3}$$

In order to implement these tests, we should know the critical values at the given significance levels, which can be obtained asymptotically through the existing limit theorems, and ensure the consistency of the tests. We can infer from the result of Csörgő and Horváth [5, p. 180], that if $E|\epsilon_1|^v < \infty$ for some $v > 2$ and

$$\limsup_{n\to\infty} n^{-\lambda} h^{-1/2} \left(\log h^{-1}\right)^{1/2} < \infty, \tag{4}$$

with $\lambda = 2^{-1} - v^{-1}$, under H_0,

$$\begin{aligned} &P\left\{A\left(\frac{n}{[nh]}\right)\sqrt{[nh]}\max_{0\leqslant k\leqslant n-[nh]}\left|\hat{\mu}_k - \hat{\mu}\right| \leqslant x + D\left(\frac{n}{[nh]}\right)\right\} \\ &= P\left\{A\left(\frac{n}{[nh]}\right)\frac{1}{\sqrt{[nh]}}\max_{0\leqslant k\leqslant n-[nh]}\left|\sum_{i=1}^{k+[nh]}\epsilon_i - \sum_{i=1}^{k}\epsilon_i - \frac{[nh]}{n}\sum_{i=1}^{n}\epsilon_i\right| \leqslant x + D\left(\frac{n}{[nh]}\right)\right\} \\ &\to \exp\left(-2e^{-x}\right) \quad \text{for all } x \in \mathbb{R}, \end{aligned} \tag{5}$$

where

$$A(x) = \sqrt{2\log x}, \tag{6}$$

$$D(x) = 2\log x + \tfrac{1}{2}\log\log x - \tfrac{1}{2}\log \pi. \tag{7}$$

Therefore, we reject H_0 if

$$M_n^* := A\left(\frac{n}{[nh]}\right)\sqrt{[nh]}\max_{0\leqslant k\leqslant n-[nh]}\left|\hat{\mu}_k - \hat{\mu}\right| - D\left(\frac{n}{[nh]}\right)$$

is large. The critical value for M_n^* is calculated from the formula in (5).

Meanwhile, with regard to the second test statistic in (3), we can see by utilizing a conventional martingale central limit theorem (CLT) that under H_0, if $E|\epsilon_1|^4 < \infty$,

$$\sqrt{\frac{3[nh]}{4n}}\left\{\sum_{k=0}^{n-[nh]}\left(\hat{\mu}_k - \hat{\mu}\right)^2 - \frac{n-[nh]+1}{[nh]}\right\} \xrightarrow{\mathrm{d}} N(0,1) \tag{8}$$

(cf. Lemma 5.4 given below). Hence, we reject H_0 if

$$Q_n^* := \sqrt{\frac{3[nh]}{4n}}\left\{\sum_{k=0}^{n-[nh]}\left(\hat{\mu}_k - \hat{\mu}\right)^2 - \frac{n-[nh]+1}{[nh]}\right\} \geqslant z_\alpha,$$

given a nominal level α, where z_α is the $(1-\alpha)$th quantile of $N(0,1)$.

Since (4) implies $n^{-1/2}h^{-1/2}(\log h^{-1})^{1/2} = O\left(n^{-1/v}\right) = o(1)$ as $n \to \infty$, M_n^* diverges to ∞ under the alternative hypothesis under which the observations have either multiple mean changes or smooth mean changes (see (ii) of Theorem 3.1 below). Similarly, by conditions (i) and (ii) of Theorem 3.2, we can see that Q_n^* is a consistent test in the cases of multiple mean changes and smooth mean changes.

Thus far, we have seen that the ME test with a varying h in an i.i.d. sample is consistent and its critical values can be obtained asymptotically. In fact, the above-mentioned results can be extended to the case of time series models, the task of which is more demanding than in the case of an i.i.d. sample.

3. ME test in time series models

3.1. ME test for a mean change in strong mixing processes

Before we proceed to analyze a general parameter case, we first consider the ME test for a mean change in the location model

$$y_i = \mu_i + \epsilon_i, \quad i \in \mathbb{Z},$$

where $\{\epsilon_i\}$ is a stationary strong mixing process with zero mean and finite $\nu(>2)$th moment. Further, we assume that

$$\alpha(n) \ll n^{-(1+\epsilon)(1+2/(\nu-2))}, \quad \epsilon > 0, \tag{9}$$

where $\alpha(n)$ denotes the strong mixing coefficient of order n. This includes invertible stationary ARMA(p, q) processes with innovations having an absolutely continuous distribution with a density $f(x)$ such that

$$\int |f(x) - f(x+y)|\,\mathrm{d}x \leqslant C|y|, \quad C > 0$$

(cf. [17,3]).

In this study, we are interested in testing the null hypothesis under which the mean is a constant over time ($\mu_1 = \cdots = \mu_n$) against the alternative hypothesis under which there exists a change in the mean.

Let $\hat{\mu}_k = [nh]^{-1}\sum_{i=k+1}^{k+[nh]} y_i$, $k = 0, 1, \ldots, n-[nh]$ where h satisfies (1), and let $\hat{\mu} = n^{-1}\sum_{i=1}^{n} y_i$. Then under the null hypothesis, $\hat{\mu}_k$ satisfies

$$\hat{\mu}_k = \mu_0 + \frac{1}{[nh]}\sum_{i=k+1}^{k+[nh]} \epsilon_i$$

for some μ_0. In addition, since $\{\epsilon_i\}$ satisfies condition (9), $\sigma^2 = \sum_{k=-\infty}^{\infty} E(\epsilon_0\epsilon_k)$ converges absolutely and there exists a probability space such that

$$\sup_{0\leqslant s\leqslant 1}\left|\sum_{i=1}^{[ns]} \epsilon_i - B(ns)\right| = o\left(n^{1/2-\lambda}\right) \quad \text{a.s.} \tag{10}$$

for some $0 < \lambda < \frac{1}{2}$ depending only on ϵ and ν, where $\{B(t), 0 \leqslant t < \infty\}$ denotes a one-dimensional Brownian motion with $EB(t)^2 = t\sigma^2$ (cf. [11]).

From (10), it is noteworthy that

$$\sqrt{n}\left(\hat{\mu} - \mu_0\right) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} \epsilon_i \xrightarrow{\mathrm{d}} N(0, \sigma^2).$$

Therefore, if $\sigma^2 > 0$ and h satisfies (4) with λ in (10), then under the null hypothesis, we have

$$M_n(\sigma) := A\left(\frac{n}{[nh]}\right)\frac{\sqrt{[nh]}}{\sigma}\max_{0\leqslant k\leqslant n-[nh]}\left|\hat{\mu}_k-\hat{\mu}\right| - D\left(\frac{n}{[nh]}\right) \xrightarrow{\mathrm{d}} \Lambda,$$

where Λ denotes a r.v. with distribution $P(\Lambda\leqslant x) = \exp\left(-2e^{-x}\right)$ for all $x \in \mathbb{R}$ (cf. Theorem 3.3 below).

Meanwhile, if $\sigma^2 > 0$ and h satisfies

$$\limsup_{n\to\infty} h^{-1}n^{-\lambda} < \infty \tag{11}$$

for λ in (10), then we obtain the following asymptotic result:

$$Q_n(\sigma) := \sqrt{\frac{3[nh]}{4n}}\left\{\frac{1}{\sigma^2}\sum_{k=0}^{n-[nh]}\left(\hat{\mu}_k-\hat{\mu}\right)^2 - \frac{n-[nh]+1}{[nh]}\right\} \xrightarrow{\mathrm{d}} N(0,1)$$

under the null hypothesis (cf. Theorem 3.4 below).

In order to study the power performance, we now consider the case where the mean varies over time, viz.,

$$\mu_i = \mu_0 + n^{-\delta}g(i/n), \tag{12}$$

where $g : [0,1] \to \mathbb{R}$ is either a non-constant step function or a continuous function. The following theorems are related to the consistency of $M_n(\sigma)$ and $Q_n(\sigma)$ and their proofs are provided in Section 5.

Theorem 3.1. *Assume that* $\sigma^2 > 0$, (12) *holds, and h satisfies* (4) *with* λ *in* (10).

(i) *If* $\delta \geqslant \frac{1}{2}$, *then*

$$M_n(\sigma) \xrightarrow{\mathrm{d}} \Lambda.$$

(ii) *If* $\delta < \frac{1}{2}$ *and* $n^{\delta-1/2}h^{-1/2}(\log h^{-1})^{1/2} = o(1)$ *as* $n \to \infty$, *then*

$$\frac{n^{\delta}}{A(n/[nh])\sqrt{[nh]}}M_n(\sigma) \xrightarrow{\mathrm{P}} \frac{1}{\sigma}\max_{0\leqslant x\leqslant 1}\left|g(x) - \int_0^1 g(x)\,\mathrm{d}x\right| > 0.$$

Theorem 3.2. *Assume that* $\sigma^2 > 0$, (12) *holds, and h satisfies* (11).

(i) *If* $\delta \geqslant \frac{1}{2}$, *then*

$$Q_n(\sigma) \xrightarrow{\mathrm{d}} N(0,1).$$

(ii) *If* $\delta < \frac{1}{2}$ *and* $n^{2\delta-1}h^{-1/2} = o(1)$ *as* $n \to \infty$, *then*

$$\sqrt{\frac{4n}{3[nh]}}n^{2\delta-1}Q_n(\sigma) \xrightarrow{\mathrm{P}} \frac{1}{\sigma^2}\int_0^1\left(g(x) - \int_0^1 g(x)\,\mathrm{d}x\right)^2\mathrm{d}x > 0.$$

Remark. An example of h is $(\log n)^{-2}$, which satisfies all the conditions for h with any λ and δ. In this case, the second statements of Theorems 3.1 and 3.2 indicate that both $M_n(\sigma)$ and $Q_n(\sigma)$ are consistent with the alternative hypothesis in (12) with $\delta < \frac{1}{2}$.

3.2. ME test for a change of general parameter

In this subsection, we deal with a general stationary time series case that includes the case presented in the previous section. Let us consider the time series $\{y_i, i \in \mathbb{Z}\}$ and let $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_d)'$ be the parameter vector that will be examined for constancy, e.g., of the mean, variance, and autocovariances, etc. In this case, our objective is to test

H_0: $\boldsymbol{\theta}$ does not change for $y_1, \ldots, y_n$ vs.
H_1: not H_0.

Similar to Section 3.1, we denote the MV of $\boldsymbol{\theta}$ based on $y_{k+1}, \ldots, y_{k+[nh]}$ with h satisfying (1) by $\hat{\boldsymbol{\theta}}_k = (\hat{\theta}_{k1}, \ldots, \hat{\theta}_{kd})$, and the full sample estimate by $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \ldots, \hat{\theta}_d)$. We now intend to test the hypotheses based on the differences $\hat{\boldsymbol{\theta}}_k - \hat{\boldsymbol{\theta}}$, $k = 0, 1, \ldots, n - [nh]$.

Suppose that under H_0,

$$\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}_0 + O_P(1/\sqrt{n}) \tag{13}$$

for some θ_0, and that $\hat{\boldsymbol{\theta}}_k$ satisfies

$$\hat{\boldsymbol{\theta}}_k = \boldsymbol{\theta}_0 + \frac{1}{[nh]} \sum_{i=k+1}^{k+[nh]} \boldsymbol{\eta}_i + \boldsymbol{\Delta}_k, \tag{14}$$

where $\{\boldsymbol{\eta}_n, n \geqslant 1\}$ is a strictly stationary sequence of $\mathbb{R}^d$-valued random vectors with zero mean and a finite vth moment for some $v > 2$. Further, we assume that we can redefine the sequence $\{\boldsymbol{\eta}_n, n \geqslant 1\}$ on a probability space together with a d-dimensional standard Brownian motion $\{\boldsymbol{W}(t), 0 \leqslant t < \infty\}$ such that

$$\sup_{0 \leqslant s \leqslant 1} \left\| \Sigma^{-1/2} \sum_{i=1}^{[ns]} \boldsymbol{\eta}_i - \boldsymbol{W}(ns) \right\| = o\left(n^{1/2-\lambda}\right) \quad \text{a.s.} \tag{15}$$

for some $0 < \lambda < \frac{1}{2}$ and positive definite (symmetric) matrix Σ of constants. Generally, $\{\boldsymbol{\eta}_n, n \geqslant 1\}$ is unobservable. However, in time series models, $\{\boldsymbol{\eta}_n, n \geqslant 1\}$ usually forms a sequence of stationary martingale differences (cf. [12]) or satisfies a strong mixing condition (cf. Section 3.1). Therefore, a broad class of time series models satisfy (15) under mild conditions (see [6,11]). In particular, if $\{\boldsymbol{\eta}_n, n \geqslant 1\}$ is a sequence of i.i.d. random vectors, then $\Sigma = E\boldsymbol{\eta}_1\boldsymbol{\eta}_1'$ and $\lambda = \frac{1}{2} - 1/v$. Condition (15) is crucial for verifying the asymptotic distribution of the test statistics presented below.

In the remaining part of this paper, $\|\boldsymbol{c}\| = \|\boldsymbol{c}\|_{\max} := \max_{1 \leqslant i \leqslant d} |c_i|$ and $\|\boldsymbol{c}\|_2 := \sqrt{\sum_{i=1}^d c_i^2}$ for $\boldsymbol{c} = (c_1, \ldots, c_d)' \in \mathbb{R}^d$. The main results of this section are given below. They are analogous to (5) and (8) in the i.i.d. sample case, and their proofs are provided in Section 5.

Theorem 3.3. *Suppose that* H_0 *and* (13)–(15) *hold. Further, assume that h satisfies* (1) *and* (4) *with* $\lambda > 0$ *in* (15). *Then if*

$$\max_{0 \leqslant k \leqslant n-[nh]} \|\boldsymbol{\Delta}_k\| = o_P\left(\frac{1}{\sqrt{nh \log h^{-1}}}\right) \quad \text{as } n \to \infty, \tag{16}$$

we have

$$P\left\{A\left(\frac{n}{[nh]}\right)\sqrt{[nh]}\max_{0\leqslant k\leqslant n-[nh]}\left\|\Sigma^{-1/2}\left(\hat{\boldsymbol{\theta}}_k-\hat{\boldsymbol{\theta}}\right)\right\|-D\left(\frac{n}{[nh]}\right)\leqslant x\right\}\rightarrow\exp\left(-2de^{-x}\right)$$

and

$$P\left\{U\left(\frac{n}{[nh]};d\right)\sqrt{[nh]}\max_{0\leqslant k\leqslant n-[nh]}\left\|\Sigma^{-1/2}\left(\hat{\boldsymbol{\theta}}_k-\hat{\boldsymbol{\theta}}\right)\right\|_2-U^2\left(\frac{n}{[nh]};d\right)\leqslant x\right\}$$
$$\rightarrow\exp\left(-e^{-x}\right)$$

for all $x\in\mathbb{R}$ *as* $n\rightarrow\infty$, *where* $A(\cdot)$ *and* $D(\cdot)$ *are defined in* (6) *and* (7), $U(x;d)=(2\log x+\mathrm{d}\log\log x+2\log 2-2\log\Gamma(d/2))^{1/2}$, *and* $\Gamma(\cdot)$ *is the gamma function.*

Theorem 3.4. *Suppose that* H_0 *and* (13)–(15) *hold. Further, assume that h satisfies* (1) *and* (11) *for* $\lambda>0$ *in* (15). *Then if*

$$\sum_{k=0}^{n-[nh]}\boldsymbol{\Delta}_k'\boldsymbol{\Delta}_k=o_P(1)\quad as\ n\rightarrow\infty, \tag{17}$$

we have

$$\sqrt{\frac{3[nh]}{4dn}}\left(\sum_{k=0}^{n-[nh]}\left\|\Sigma^{-1/2}\left(\hat{\boldsymbol{\theta}}_k-\hat{\boldsymbol{\theta}}\right)\right\|_2^2-\frac{(n-[nh]+1)d}{[nh]}\right)\xrightarrow{\mathrm{d}}N(0,1). \tag{18}$$

In view of the results of Theorems 3.3 and 3.4, we can construct the test statistics

$$M_{n1}:=A\left(\frac{n}{[nh]}\right)\sqrt{[nh]}\max_{0\leqslant k\leqslant n-[nh]}\left\|\hat{\Sigma}^{-1/2}\left(\hat{\boldsymbol{\theta}}_k-\hat{\boldsymbol{\theta}}\right)\right\|-D\left(\frac{n}{[nh]}\right), \tag{19}$$

$$M_{n2}:=U\left(\frac{n}{[nh]};d\right)\left(\sqrt{[nh]}\max_{0\leqslant k\leqslant n-[nh]}\left\|\hat{\Sigma}^{-1/2}\left(\hat{\boldsymbol{\theta}}_k-\hat{\boldsymbol{\theta}}\right)\right\|_2-U\left(\frac{n}{[nh]};d\right)\right), \tag{20}$$

$$Q_n:=\sqrt{\frac{3[nh]}{4dn}}\left(\sum_{k=0}^{n-[nh]}\left\|\hat{\Sigma}^{-1/2}\left(\hat{\boldsymbol{\theta}}_k-\hat{\boldsymbol{\theta}}\right)\right\|_2^2-\frac{(n-[nh]+1)d}{[nh]}\right), \tag{21}$$

where $\hat{\Sigma}$ is a consistent estimator of Σ. We reject H_0 at the level α if M_{n1}, M_{n2} and Q_n are larger than $c_{\alpha1}$, $c_{\alpha2}$ and z_α, respectively, where $c_{\alpha1}=-\log\left(-\frac{1}{2d}\log(1-\alpha)\right)$, $c_{\alpha2}=-\log(-\log(1-\alpha))$ and z_α is the $(1-\alpha)$th quantile of $N(0,1)$. The performance of the tests will be evaluated through a simulation study reported in Section 4.

As we have observed in the linear process case, we may verify the consistency of the tests for certain types of alternatives. However, instead of going into details on this issue, we consider the method to allocate the locations of abrupt changes.

Providing a change that occurs at $[n\tau]$ for $\tau\in(0,1)$, we employ

$$\hat{\tau}_n=\min\left\{\frac{k}{n}:[nh]\leqslant k\leqslant n-[nh]\text{ and }\left\|\hat{\boldsymbol{\theta}}_k-\hat{\boldsymbol{\theta}}_{k-[nh]}\right\|=\max_{[nh]\leqslant k\leqslant n-[nh]}\left\|\hat{\boldsymbol{\theta}}_k-\hat{\boldsymbol{\theta}}_{k-[nh]}\right\|\right\}$$

as an estimate of τ. It is because $\hat{\tau}$ is weakly consistent under the alternative hypothesis with a single change:

$\mathrm{H_a}$: $\boldsymbol{\theta}$ changes at $[n\tau]$ from $\boldsymbol{\theta}_1$ to $\boldsymbol{\theta}_2$, where $0 < \tau < 1$ and $\boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_2$ are unknown.

The result is summarized as follows.

Theorem 3.5. *Suppose that* $\mathrm{H_a}$ *holds and that under* $\mathrm{H_a}$, $\hat{\boldsymbol{\theta}}_k$ *satisfies*

$$\hat{\boldsymbol{\theta}}_k = \begin{cases} \boldsymbol{\theta}_1 + \frac{1}{[nh]} \sum_{i=k+1}^{k+[nh]} \boldsymbol{\eta}_{1k} + \boldsymbol{\Delta}_{1k}, & k = 0, 1, \ldots, [n\tau] - [nh] - 1, \\ \boldsymbol{\theta}_2 + \frac{1}{[nh]} \sum_{i=k+1}^{k+[nh]} \boldsymbol{\eta}_{2k} + \boldsymbol{\Delta}_{2k}, & k = [n\tau], [n\tau] + 1, \ldots, n - [nh], \end{cases}$$

where $\{\boldsymbol{\eta}_{in}, n \geqslant 1\}$ *satisfies* (15) *for some positive definite matrix* Σ_i *and* $0 < \lambda_i < \frac{1}{2}$, $i = 1, 2$. *If* (4) *with* $\lambda = \lambda_1 \wedge \lambda_2$ *holds and*

$$\max_k \|\boldsymbol{\Delta}_{ik}\| = o_P(1) \tag{22}$$

for $i = 1, 2$, *then*

$$\left|\hat{\tau} - \tau\right| = O_P(h). \tag{23}$$

4. Simulation results

In this section, we evaluate the tests proposed in Section 3 based on the location model:

$$y_i = \mu_i + \epsilon_i, \quad i = 1, 2, \ldots, n.$$

To achieve this task, we consider the null hypothesis

$$\mathcal{K}_0 : \mu_i = 0 \quad \text{for all } i = 1, 2, \ldots, n, \tag{24}$$

and the three alternative hypotheses

$$\mathcal{K}_1 : \mu_i = \begin{cases} 0, & i = 1, \ldots, [n\tau], \\ 1, & i = [n\tau] + 1, \ldots, n, \end{cases} \tag{25}$$

$$\mathcal{K}_2 : \mu_i = \begin{cases} 0, & i = 1, \ldots, [n\tau_1], \\ 1, & i = [n\tau_1] + 1, \ldots, [n\tau_2], \\ 0, & i = [n\tau_2] + 1, \ldots, n, \end{cases} \tag{26}$$

$$\mathcal{K}_3 : \mu_i = (1/2) * (-1)^{[p*i/n]}, \quad i = 1, \ldots, n. \tag{27}$$

Note that $\mathcal{K}_i$, $i = 1, 2, 3$, represent one single change, a temporary shift, and a periodic change, respectively. The tests are compared with the fixed-h ME test (ME_h) with $h = 0.1, 0.2$ and 0.5 (cf. [4]) and the cusum test T_n proposed by Lee et al. [12]. In this simulation, we use the data sets with the sample size $n = 1000$, 2000 and 3000, and employ the bandwidth $h = h_n = (\log n)^{-2}$ (cf. Remark in Section 3.1). We perform a test at the nominal level $\alpha = 0.1$ by using the critical values in Table 1 (cf. [4,12]). The empirical sizes and powers are calculated as a proportion of the rejection number of the null hypothesis out of 1000 repetitions.

First, we evaluate the performance of the tests in the i.i.d. sample. To achieve this task, we generate i.i.d. $N(0, 1)$ r.v.'s ϵ_i, and use the sample variance of residuals $\hat{\epsilon}_k = y_{k+[([nh]+1)/2]} - \hat{\mu}_k$,

Table 1
Critical values

$ME_{0.1}$	$ME_{0.2}$	$ME_{0.5}$	Q_n	M_{n1}	M_{n2}	T_n
0.981	1.217	1.356	1.282	2.944	2.250	1.488

Table 2
Empirical sizes under $\mathcal{K}_0$ for the i.i.d. sample

n	$ME_{0.1}$	$ME_{0.2}$	$ME_{0.5}$	Q_n	M_{n1}	M_{n2}	T_n
1000	0.121	0.123	0.114	0.148	0.051	0.067	0.107
2000	0.136	0.126	0.129	0.116	0.050	0.069	0.109
3000	0.120	0.107	0.125	0.105	0.051	0.073	0.097

Table 3
Empirical powers under $\mathcal{K}_1$ for the i.i.d. sample

τ	n	$ME_{0.1}$	$ME_{0.2}$	$ME_{0.5}$	Q_n	M_{n1}	M_{n2}	T_n
0.5	1000	1	1	1	1	1	1	1
	2000	1	1	1	1	1	1	1
	3000	1	1	1	1	1	1	1
0.95	1000	0.958	0.715	0.700	0.986	0.991	0.993	0.990
	2000	1	0.964	0.943	1	1	1	1
	3000	1	0.999	0.993	1	1	1	1
0.975	1000	0.308	0.214	0.294	0.584	0.851	0.872	0.362
	2000	0.625	0.389	0.480	0.929	0.997	0.997	0.755
	3000	0.849	0.520	0.581	0.993	1	1	0.957
0.99	1000	0.129	0.125	0.148	0.164	0.111	0.131	0.125
	2000	0.152	0.147	0.175	0.202	0.354	0.396	0.162
	3000	0.190	0.164	0.224	0.238	0.689	0.718	0.218

$k = 0, 1, \ldots, n - [nh]$ as an estimate of $\sigma^2 = Var(\epsilon_1)$. We can easily verify that this estimator is consistent under the hypotheses in (24)–(27).

The figures in Table 2 denote empirical sizes, and the result shows that in most cases, no tests produce severe size distortions. To examine the power of a test, we first consider the alternative hypothesis $\mathcal{K}_1$ in (25) with $\tau = 0.5, 0.95, 0.975$ and 0.99. Table 3 shows that the empirical powers of all tests are fairly good in most cases when τ is smaller than 0.975. Further, it shows that M_{n1} and M_{n2} produce better powers than the others when $\tau > 0.95$, and that performance improves further when $\tau = 0.99$ and $n \geqslant 2000$. Intuitively, the results appear to be designed to better detect the changes near the ending points when the sample size is fairly large.

For $\mathcal{K}_2$ in (26), we take into account of the cases in which $\tau_1 = 0.5$ and $\tau_2 = 0.6$, 0.55, 0.525. Table 4 shows that all the tests produce good powers in most cases except for the case in which $\tau_2 - \tau_1 = 0.025$, and that Q_n, M_{n1} and M_{n2} produce better powers than the others when $\tau_2 - \tau_1 = 0.025$. Further, it shows that M_{ni}'s have a tendency to decrease in power as τ_1 gets closer to τ_2, and the others experience a more severe power loss than M_{ni}'s.

For $\mathcal{K}_3$ in (27), we consider the cases of the period $(= 2p^{-1})$ 0.1, 0.2 and 0.5. Table 5 shows that Q_n, M_{n1} and M_{n2} have fairly good powers in all the cases, and that the ME_h have good

Table 4
Empirical powers under $\mathcal{K}_2$ for the i.i.d. sample

τ_2	n	$\mathrm{ME}_{0.1}$	$\mathrm{ME}_{0.2}$	$\mathrm{ME}_{0.5}$	Q_n	M_{n1}	M_{n2}	T_n
0.6	1000	1	1	0.997	1	1	1	0.998
	2000	1	1	1	1	1	1	1
	3000	1	1	1	1	1	1	1
0.55	1000	0.996	0.935	0.724	0.994	0.988	0.991	0.609
	2000	1	1	0.942	1	1	1	0.919
	3000	1	1	0.993	1	1	1	0.999
0.525	1000	0.610	0.429	0.319	0.764	0.857	0.874	0.269
	2000	0.917	0.719	0.442	0.978	1	1	0.367
	3000	0.992	0.878	0.610	0.997	1	1	0.527

Table 5
Empirical powers $\mathcal{K}_3$ for the i.i.d. sample

Period	n	$\mathrm{ME}_{0.1}$	$\mathrm{ME}_{0.2}$	$\mathrm{ME}_{0.5}$	Q_n	M_{n1}	M_{n2}	T_n
0.1	1000	0.083	0.081	0.093	1	0.997	0.998	0.532
	2000	0.101	0.101	0.102	1	1	1	0.882
	3000	0.099	0.094	0.117	1	1	1	0.992
0.2	1000	1	0.105	1	1	1	1	0.998
	2000	1	0.091	1	1	1	1	1
	3000	1	0.109	1	1	1	1	1
0.5	1000	1	1	0.097	1	1	1	1
	2000	1	1	0.126	1	1	1	1
	3000	1	1	0.112	1	1	1	1

powers except when h is proportional to the period of the mean function. In this case, our tests turned out to produce much better powers.

Now, we turn our attention to the time series models. In order to do so, we consider the case where $\{\epsilon_i\}$ follows the AR(1) model

$$\epsilon_i = \phi\epsilon_{i-1} + e_i,$$

with $\phi = -0.5$, 0.0 and 0.5. In this case, e_i's are generated from $N(0, 1)$, and 100 initial observations are discarded to remove initialization effects. To estimate $\sigma^2 = E(\epsilon_0^2) + 2\sum_{k=1}^{\infty} E(\epsilon_0\epsilon_k)$, we use the estimator

$$\hat{\sigma}^2 = \frac{1}{n - N_n + 1} \sum_{j=1}^{n-N_n+1} \left(\hat{\gamma}_j(0) + 2 \sum_{l=1}^{N_n^{1/4}} \hat{\gamma}_j(l) \right),$$

where $N_n = [n/\log n]$ and

$$\hat{\gamma}_j(l) = \frac{1}{N_n} \sum_{i=j}^{j+N_n-l-1} \left(y_i - \frac{1}{N_n} \sum_{i=j}^{j+N-1} y_i \right) \left(y_{i+l} - \frac{1}{N_n} \sum_{i=j}^{j+N-1} y_i \right)$$

Table 6
Empirical sizes under $\mathcal{K}_0$ for the AR(1) process

ϕ	n	$\mathrm{ME}_{0.1}$	$\mathrm{ME}_{0.2}$	$\mathrm{ME}_{0.5}$	Q_n	M_{n1}	M_{n2}	T_n
−0.5	1000	0.123	0.104	0.095	0.138	0.060	0.087	0.098
	2000	0.145	0.137	0.126	0.192	0.089	0.117	0.127
	3000	0.134	0.120	0.121	0.136	0.086	0.116	0.109
0	1000	0.129	0.149	0.131	0.108	0.046	0.062	0.113
	2000	0.111	0.118	0.107	0.105	0.048	0.069	0.102
	3000	0.122	0.131	0.118	0.114	0.058	0.077	0.108
0.5	1000	0.134	0.136	0.140	0.141	0.036	0.051	0.139
	2000	0.117	0.111	0.136	0.115	0.036	0.052	0.110
	3000	0.120	0.116	0.126	0.110	0.041	0.055	0.106

for $j = 0, 1, \ldots, n - N_n + 1$. This estimator is consistent under the null and alternative hypotheses in (25)–(27). In this case, we perform a test within the same framework as in the i.i.d. sample case. However, we only report the simulation result for $\mathcal{K}_1$ and $\mathcal{K}_2$ in the power study since all the tests perform unsatisfactorily in the case of $\mathcal{K}_3$ due to the poor performance of $\hat{\sigma}^2$.

Table 6 demonstrates that none of the tests produce severe size distortions. Meanwhile, Tables 7 and 8 show that the empirical powers of all the tests are fairly good in most cases when $\tau \leqslant 0.95$ and $\tau_2 - \tau_1 \geqslant 0.05$ as observed in the case of the i.i.d. sample. Further, it shows that there is a tendency to decrease in the power as ϕ increases, and M_{n1} and M_{n2} produce better powers than the others when n is large, $\tau > 0.95$, and $\tau_2 - \tau_1 < 0.05$.

From the simulation result, it can be concluded that the ME tests with time varying bandwidths perform satisfactorily. As compared to the original ME and cusum tests, our tests were found to have an advantage in that they could produce better powers in detecting the changes near ending points, the temporarily shifted changes for a short interval of time, and the periodic changes. The new ME test is believed to be an efficient functional tool in detecting changes in time series with fairly long lags.

5. Proofs

For simplicity of notation, let $A_n = A(n/[nh])$, $D_n = D(n/[nh])$, $U_n = U(n/[nh]; d)$, and let $\bar{x}_k = [nh]^{-1} \sum_{i=k+1}^{k+[nh]} x_i$ and $\bar{x} = n^{-1} \sum_{i=1}^{n} x_i$ for r.v.'s (or random vectors) $x_1, \ldots, x_n$. Further, let $g_n(k) := [nh]^{-1} \sum_{i=k+1}^{k+[nh]} g\,(i/n) - n^{-1} \sum_{i=1}^{n} g\,(i/n)$, $k = 0, 1, \ldots, n - [nh]$ for a real-valued function g on [0, 1].

Let $\{\boldsymbol{x}_n, n \geqslant 1\}$ be a strictly stationary sequence of $\mathbb{R}^d$-valued random vectors with zero mean and

$$\sup_{0 \leqslant s \leqslant 1} \left\| \Sigma^{-1/2} \sum_{i=1}^{[ns]} \boldsymbol{x}_i - \boldsymbol{W}(ns) \right\| = o\left(n^{1/2-\lambda}\right) \quad \text{a.s.} \tag{28}$$

for some non-singular matrix Σ and $0 < \lambda < \frac{1}{2}$. Before proving Theorems 3.1–3.5, we investigate some properties of $\bar{\boldsymbol{x}}_k = [nh]^{-1} \sum_{i=k+1}^{k+[nh]} \boldsymbol{x}_i$, $k = 0, 1, \ldots, n - [nh]$.

Table 7
Empirical powers under $\mathcal{K}_1$ for the AR(1) process

ϕ	τ	n	$\mathrm{ME}_{0.1}$	$\mathrm{ME}_{0.2}$	$\mathrm{ME}_{0.5}$	Q_n	M_{n1}	M_{n2}	T_n
−0.5	0.5	1000	1	1	1	1	1	1	1
		2000	1	1	1	1	1	1	1
		3000	1	1	1	1	1	1	1
	0.95	1000	1	0.942	0.903	0.999	1	1	1
		2000	1	1	0.999	1	1	1	1
		3000	1	1	1	1	1	1	1
	0.975	1000	0.590	0.339	0.432	0.824	1	1	0.718
		2000	0.959	0.707	0.726	1	1	1	0.999
		3000	0.997	0.909	0.880	1	1	1	1
	0.99	1000	0.144	0.145	0.172	0.155	0.322	0.367	0.167
		2000	0.248	0.190	0.267	0.357	0.890	0.909	0.277
		3000	0.325	0.249	0.323	0.452	0.993	0.996	0.385
0.0	0.5	1000	1	1	1	1	0.994	0.999	1
		2000	1	1	1	1	1	1	1
		3000	1	1	1	1	1	1	1
	0.95	1000	0.938	0.671	0.662	0.984	0.982	0.989	0.971
		2000	0.998	0.955	0.917	1	1	1	1
		3000	1	0.997	0.989	1	1	1	1
	0.975	1000	0.308	0.246	0.311	0.606	0.858	0.869	0.382
		2000	0.641	0.400	0.427	0.919	0.996	0.997	0.730
		3000	0.852	0.540	0.588	0.990	1	1	0.944
	0.99	1000	0.138	0.136	0.147	0.139	0.117	0.136	0.139
		2000	0.154	0.151	0.175	0.171	0.377	0.409	0.179
		3000	0.161	0.162	0.194	0.232	0.688	0.717	0.195
0.5	0.5	1000	1	1	1	1	0.506	0.582	1
		2000	1	1	1	1	0.780	0.839	1
		3000	1	1	1	1	0.958	0.974	1
	0.95	1000	0.362	0.281	0.314	0.564	0.301	0.353	0.375
		2000	0.664	0.400	0.443	0.803	0.616	0.676	0.693
		3000	0.859	0.545	0.582	0.927	0.856	0.880	0.897
	0.975	1000	0.160	0.151	0.198	0.298	0.153	0.181	0.187
		2000	0.195	0.179	0.223	0.379	0.380	0.429	0.235
		3000	0.262	0.204	0.253	0.545	0.595	0.637	0.295
	0.99	1000	0.149	0.139	0.161	0.155	0.034	0.048	0.147
		2000	0.143	0.134	0.141	0.133	0.044	0.063	0.122
		3000	0.143	0.129	0.136	0.159	0.101	0.124	0.137

Lemma 5.1. *If h satisfies* (1) *and* (4) *with λ in* (28), *then we have*

$$\lim_{n\to\infty} P\left\{ A_n\sqrt{[nh]} \max_{0\leqslant k\leqslant n-[nh]} \left\| \Sigma^{-1/2}\bar{\boldsymbol{x}}_k \right\| - D_n \leqslant x \right\} = \exp\left(-2de^{-x}\right).$$

Proof. Due to (28) and (4), we have

$$\begin{aligned} &A_n\sqrt{[nh]} \max_{0\leqslant k\leqslant n-[nh]} \left\| \Sigma^{-1/2}\bar{\boldsymbol{x}}_k \right\| - D_n \\ &\quad = A_n \frac{1}{\sqrt{[nh]}} \sup_{0\leqslant s\leqslant n-[nh]} \|\boldsymbol{W}(s+[nh]) - \boldsymbol{W}(s)\| - D_n + o_P(1). \end{aligned}$$

Table 8
Empirical powers under $\mathcal{K}_2$ for the AR(1) process

ϕ	τ_2	n	$\mathrm{ME}_{0.1}$	$\mathrm{ME}_{0.2}$	$\mathrm{ME}_{0.5}$	Q_n	M_{n1}	M_{n2}	T_n
−0.5	0.6	1000	1	1	1	1	1	1	0.999
		2000	1	1	1	1	1	1	1
		3000	1	1	1	1	1	1	1
	0.55	1000	1	0.979	0.609	0.996	1	1	0.463
		2000	1	1	0.930	1	1	1	0.913
		3000	1	1	1	1	1	1	1
	0.525	1000	0.737	0.439	0.250	0.565	0.999	1	0.178
		2000	0.995	0.805	0.364	0.990	1	1	0.252
		3000	1	0.958	0.553	1	1	1	0.451
0.0	0.6	1000	1	1	0.985	1	0.995	0.996	0.978
		2000	1	1	1	1	1	1	1
		3000	1	1	1	1	1	1	1
	0.55	1000	0.993	0.849	0.512	0.947	0.964	0.970	0.414
		2000	1	0.992	0.806	1	1	1	0.710
		3000	1	1	0.959	1	1	1	0.961
	0.525	1000	0.489	0.337	0.231	0.474	0.815	0.847	0.202
		2000	0.828	0.579	0.314	0.887	0.995	0.996	0.243
		3000	0.960	0.762	0.432	0.988	1	1	0.346
0.5	0.6	1000	0.979	0.895	0.634	0.842	0.365	0.418	0.532
		2000	1	0.996	0.889	0.989	0.752	0.805	0.828
		3000	1	1	0.976	0.999	0.956	0.967	0.977
	0.55	1000	0.546	0.406	0.265	0.507	0.234	0.284	0.240
		2000	0.835	0.610	0.382	0.748	0.561	0.614	0.321
		3000	0.964	0.832	0.545	0.929	0.829	0.850	0.456
	0.525	1000	0.227	0.209	0.180	0.234	0.128	0.156	0.164
		2000	0.327	0.248	0.191	0.352	0.330	0.375	0.166
		3000	0.440	0.320	0.245	0.527	0.576	0.613	0.189

By the scale transformation of a Wiener process and Theorem 7.2.4 of Révész [15, p. 72], we get

$$
\begin{aligned}
&P\left\{A_n \frac{1}{\sqrt{[nh]}} \sup_{0\leqslant s\leqslant n-[nh]} \|\boldsymbol{W}(s+[nh])-\boldsymbol{W}(s)\| - D_n \leqslant x\right\} \\
&= \left[P\left\{A_n \frac{1}{\sqrt{[nh]}} \sup_{0\leqslant s\leqslant n/[nh]-1} |W(s+1)-W(s)| - D_n \leqslant x\right\}\right]^d \\
&\to \exp\left(-2de^{-x}\right).
\end{aligned}
$$

Thus, the lemma is established. □

Lemma 5.2. *Under the assumptions of Lemma* 5.1, *we have*

$$
\lim_{n\to\infty} P\left\{U_n\left(\sqrt{[nh]} \max_{0\leqslant k\leqslant n-[nh]} \left\|\Sigma^{-1/2}\bar{\boldsymbol{x}}_k\right\| - U_n\right) \leqslant x\right\} = \exp\left(-e^{-x}\right).
$$

Proof. Since (28) and (4) hold, we can obtain the result by applying Theorem 10 of Albin [1]. $\square$

Lemma 5.3. *Under the assumptions of Lemma* 5.1, *we have*

$$\lim_{n\to\infty} P\left\{A_n\sqrt{\frac{[nh]}{2}}\max_{0\leqslant k\leqslant n-[nh]}\left\|\Sigma^{-1/2}\left(\bar{\boldsymbol{x}}_k-\bar{\boldsymbol{x}}_{k-[nh]}\right)\right\|-D_n\leqslant x\right\}=\exp\left(-2de^{-x}\right).$$

Proof. Since (28) and (4) hold, we can obtain the result by applying Theorem A.1 of Bickel and Rosenblatt [2]. $\square$

Lemma 5.4. *Let* $z_1,\ldots,z_n$ *be i.i.d. random vectors with zero mean*, *covariance matrix* I_d *and a finite fourth moment*, *where* I_d *is the* d*-dimensional identity matrix. Then we have*

$$\sqrt{\frac{3[nh]}{4nd}}\left(\sum_{k=0}^{n-[nh]}\|\bar{z}_k\|_2^2-\frac{(n-[nh]+1)d}{[nh]}\right)\xrightarrow{\mathrm{d}} N(0,1). \tag{29}$$

Proof. Write

$$\sum_{k=0}^{n-[nh]}\|\bar{z}_k\|_2^2=I_n^*+2II_n^*,$$

where

$$I_n^*=\sum_{i=1}^{n}\left(\sum_{k=0}^{n-[nh]}[nh]^{-2}I(i-[nh]\leqslant k<i)\right)z_i'z_i$$

and

$$II_n^*=\sum_{j=2}^{n}\sum_{i=1}^{j-1}\left(\sum_{k=0}^{n-[nh]}[nh]^{-2}I(j-[nh]\leqslant k<i)\right)z_i'z_j.$$

Since $z_1'z_1,\ldots,z_n'z_n$ are i.i.d. random variables with mean d and variance $Var(z_1'z_1)<\infty$, we have

$$\frac{I_n^*-EI_n^*}{\sqrt{Var(I_n^*)}}=\frac{I_n^*-\dfrac{(n-[nh]+1)d}{[nh]}}{\sqrt{\left(2\sum_{i=1}^{[nh]-1}\dfrac{i^2}{[nh]^4}+\dfrac{n-2[nh]+2}{[nh]^2}\right)Var(z_1'z_1)}}\xrightarrow{\mathrm{d}} N(0,1)$$

by using the Lindeberg theorem. Hence,

$$I_n^*=\frac{(n-[nh]+1)d}{[nh]}+O_P\left(\frac{1}{\sqrt{nh^2}}\right) \tag{30}$$

as $n\to\infty$.

Next, if we set $\xi_{nj} = \sum_{i=1}^{j-1} \left(\sum_{k=0}^{n-[nh]} [nh]^{-2} I(j - [nh] \leqslant k < i) \right) z_i' z_j$ and $\mathcal{F}_j = \sigma(z_1, \ldots, z_j)$, then $\{\xi_{nj}, \mathcal{F}_j\}$ forms a square integrable martingale difference array. Let $\sigma_{nj}^2 = E\left(\xi_{nj}^2 | \mathcal{F}_{j-1}\right)$ and $s_n^2 = nd/3[nh]$. Then it holds that

$$s_n^{-1} II_n^* = s_n^{-1} \sum_{j=2}^{n} \xi_{nj} \xrightarrow{\mathrm{d}} N(0,1), \tag{31}$$

if the following conditions are satisfied (cf. [7, p. 307]):

$$s_n^{-2} \sum_{j=2}^{n} \sigma_{nj}^2 \xrightarrow{\mathrm{P}} 1, \tag{32}$$

$$\forall \epsilon > 0, \quad s_n^{-2} \sum_{j=2}^{n} E\left(\xi_{nj}^2 I\left(|\xi_{nj}| \geqslant \epsilon s_n\right)\right) \to 0. \tag{33}$$

Since (30) and (31) imply (29), it is enough to show that (32) and (33) hold.

In order to show (32), write

$$s_n^{-2} \sum_{j=2}^{n} \sigma_{nj}^2 - 1 = s_n^{-2} \left(\sum_{j=2}^{n} \sigma_{nj}^2 - \sum_{j=[nh]}^{n-[nh]} \sigma_{nj}^2 \right) + s_n^{-2} \left(\sum_{j=[nh]}^{n-[nh]} \sigma_{nj}^2 - s_n^2 \right). \tag{34}$$

Since

$$E\left| \sum_{j=2}^{n} \sigma_{nj}^2 - \sum_{j=[nh]}^{n-[nh]+2} \sigma_{nj}^2 \right| \leqslant 2[nh] \max_{2 \leqslant j \leqslant n} E\sigma_{nj}^2 \leqslant \frac{1}{[nh]^2} \sum_{i=j-[nh]+1}^{j-1} E(z_i' z_i) \leqslant \frac{d}{[nh]},$$

the first term of the right-hand side of (34) converges to 0 in L_1.

Note that $\sigma_{nj}^2 \stackrel{\mathrm{D}}{=} \sigma_{n,[nh]}^2$ for all j with $[nh] \leqslant j \leqslant n - [nh] + 2$ and note that σ_{nj}^2 and $\sigma_{nj'}^2$ are independent for $|j - j'| > [nh] - 2$. Hence,

$$\begin{aligned}
&E\left(\sum_{j=[nh]}^{n-[nh]+2} \sigma_{nj}^2 - s_n^2 \right)^2 \\
&= Var\left(\sum_{j=[nh]}^{n-[nh]+2} \sigma_{nj}^2 \right) + \left(\sum_{j=[nh]}^{n-[nh]+2} E\sigma_{nj}^2 - s_n^2 \right)^2 \\
&\leqslant 2n[nh] Var\left(\sigma_{n,[nh]}^2\right) + \left(\frac{(n - 2[nh] + 3)d}{[nh]^4} \sum_{i=1}^{[nh]-1} i^2 - s_n^2 \right)^2 \\
&= O\left(n/[nh]\right) + O(1) + O(n^2/[nh]^4) \\
&= o(s_n^4)
\end{aligned}$$

by the Cauchy–Schwarz inequality, because

$$Var\left(\sigma_{n,[nh]}^2\right) \leqslant E\left(\sigma_{n,[nh]}^4\right) \leqslant \max_j E\left(\xi_{nj}^4\right) \leqslant C_1/[nh]^2 \tag{35}$$

for some positive real value C_1. Therefore, the second term of the right-hand side of (34) converges to 0 in L_2. So, the left-hand side of (34) converges to 0 in probability, and thus (32) holds.

For (33), note that given $\epsilon > 0$,

$$\frac{1}{s_n^2}\sum_{j=2}^{n} E\left(\xi_{nj}^2 I\left(\left|\xi_{nj}\right| > \epsilon s_n\right)\right) \leqslant \frac{1}{s_n^2}\sum_{j=2}^{n}\frac{E\left(\xi_{nj}^4\right)}{\epsilon^2 s_n^2} \leqslant \frac{C_1 n}{\epsilon^2 s_n^4 [nh]^2} = \frac{9C_1}{\epsilon^2 d^2 n} \to 0$$

by (35). This completes the proof. □

Lemma 5.5. *If h satisfies* (1) *and* (11) *with λ in* (28), *then we have*

$$\sqrt{\frac{3[nh]}{4nd}}\left(\sum_{k=0}^{n-[nh]}\left\|\Sigma^{-1/2}\bar{\boldsymbol{x}}_k\right\|_2^2 - \frac{(n-[nh]+1)d}{[nh]}\right) \xrightarrow{\text{d}} N(0,1).$$

Proof. Let $\boldsymbol{w}_i = \boldsymbol{W}(i) - \boldsymbol{W}(i-1)$ for $i = 1, 2, \ldots, n$. Since $\boldsymbol{w}_i$ are i.i.d. $N(\boldsymbol{0}, I_d)$, by Lemma 5.4

$$\sqrt{\frac{3[nh]}{4nd}}\left(\sum_{k=0}^{n-[nh]}\|\bar{\boldsymbol{w}}_k\|_2^2 - \frac{(n-[nh]+1)d}{[nh]}\right) \xrightarrow{\text{d}} N(0,1).$$

Due to (28) and (11), we have

$$\begin{aligned}\sum_{k=0}^{n-[nh]}\left\|\Sigma^{-1/2}\bar{\boldsymbol{x}}_k - \bar{\boldsymbol{w}}_k\right\|_2^2 &\leqslant \left(\frac{2d\sqrt{n}}{[nh]}\sup_{0\leqslant s\leqslant 1}\left\|\Sigma^{-1/2}\sum_{i=1}^{[ns]}\boldsymbol{x}_i - \boldsymbol{W}([ns])\right\|\right)^2 = o_P\left(n^{-2\lambda}h^{-2}\right)\\ &= o_P(1).\end{aligned}$$

Therefore, we can obtain that

$$\begin{aligned}&\left|\sum_{k=0}^{n-[nh]}\left\|\Sigma^{-1/2}\bar{\boldsymbol{x}}_k\right\|_2^2 - \sum_{k=0}^{n-[nh]}\left\|\Sigma^{-1/2}\bar{\boldsymbol{w}}_k\right\|_2^2\right|\\ &\quad\leqslant \sum_{k=0}^{n-[nh]}\left\|\Sigma^{-1/2}\bar{\boldsymbol{x}}_k - \bar{\boldsymbol{w}}_k\right\|_2^2 + 2\left(\sum_{k=0}^{n-[nh]}\left\|\Sigma^{-1/2}\bar{\boldsymbol{x}}_k - \bar{\boldsymbol{w}}_k\right\|_2^2\sum_{k=0}^{n-[nh]}\|\bar{\boldsymbol{w}}_k\|_2^2\right)^{1/2}\\ &\quad= o_P\left(n^{-2\lambda}h^{-2}\right) + \left\{o_P\left(n^{-2\lambda}h^{-2}\right)\left(O_P\left(h^{-1/2}\right) + O\left(h^{-1}\right)\right)\right\}^{1/2}\\ &\quad= o_P\left(h^{-1/2}\right)\end{aligned}$$

as $n \to \infty$. This completes the proof. □

Proof of Theorem 3.1.

(i) By the triangle inequality, for $\delta \geqslant \frac{1}{2}$,

$$\begin{aligned}&\left|M_n(\sigma) - A_n\sqrt{[nh]}\sigma^{-1}\max_{0\leqslant k\leqslant n-[nh]}|\bar{\epsilon}_k| + D_n\right|\\ &\quad\leqslant A_n\sqrt{[nh]}\sigma^{-1}\left(n^{-\delta}\max_{0\leqslant k\leqslant n-[nh]}|g_n(k)| + |\bar{\epsilon}|\right)\\ &\quad= O\left(n^{1/2}h^{1/2}(\log h^{-1})^{1/2}\right)\left(O\left(n^{-\delta}\right) + O_P\left(n^{-1/2}\right)\right)\\ &\quad= o_P(1).\end{aligned}$$

Hence, by Lemma 5.1, $M_n(\sigma)$ converges to Λ in distribution.

(ii) By the triangle inequality and Lemma 5.1,

$$\begin{aligned}
&\left|\frac{n^{\delta}}{A_n\sqrt{[nh]}}M_n(\sigma)-\sigma^{-1}\max_{0\leqslant k\leqslant n-[nh]}|g_n(k)|\right|\\
&\quad\leqslant\frac{n^{\delta}}{A_n\sqrt{[nh]}}\left(A_n\sqrt{[nh]}\sigma^{-1}\max_{0\leqslant k\leqslant n-[nh]}|\bar{\epsilon}_k|-D_n\right)+2\frac{n^{\delta}D_n}{A_n\sqrt{[nh]}}+\sigma^{-1}|\bar{\epsilon}|\\
&\quad=O_P\left(n^{\delta-1/2}h^{-1/2}(\log h^{-1})^{-1/2}\right)+O\left(n^{\delta-1/2}h^{-1/2}(\log h^{-1})^{1/2}\right)+O_P\left(n^{-1/2}\right).
\end{aligned}$$

Since $n^{\delta-1/2}h^{-1/2}(\log h^{-1})^{1/2}\to 0$ and $\max_k|g_n(k)|\to\max_x\left|g(x)-\int_0^1 g(x)\,\mathrm{d}x\right|$ as $n\to\infty$, (ii) is asserted. $\square$

Proof of Theorem 3.2. First, note that

$$\begin{aligned}
Q_n(\sigma)&=\sqrt{\frac{3[nh]}{4n}}\left(\frac{1}{\sigma^2}\sum_{k=0}^{n-[nh]}\bar{\epsilon}_k^2-\frac{n-[nh]+1}{[nh]}\right)\\
&\quad+\sqrt{\frac{3[nh]}{4n}}n^{-2\delta}\frac{1}{\sigma^2}\sum_{k=0}^{n-[nh]}g_n^2(k)\\
&\quad+\sqrt{\frac{3[nh]}{4n}}\frac{1}{\sigma^2}\sum_{k=0}^{n-[nh]}\left(\bar{\epsilon}^2+2\bar{\epsilon}\bar{\epsilon}_k+2n^{-\delta}g_n(k)\bar{\epsilon}_k-2n^{-\delta}g_n(k)\bar{\epsilon}\right)\\
&=:I_n+II_n+III_n,
\end{aligned}$$

and that I_n converges to $N(0,1)$ in distribution. In addition, note that

$$\frac{1}{n}\sum_{k=0}^{n-[nh]}g_n^2(k)\to\int_0^1\left(g(x)-\int_0^1 g(x)\,\mathrm{d}x\right)^2\mathrm{d}x, \tag{36}$$

$$0\leqslant\sum_{k=0}^{n-[nh]}\left(\bar{\epsilon}^2+2\bar{\epsilon}\bar{\epsilon}_k\right)\leqslant 3\left(\sqrt{n}\bar{\epsilon}\right)^2+o_P(1)=O_P(1), \tag{37}$$

and

$$E\left(\frac{1}{\sqrt{n}}\sum_{k=0}^{n-[nh]}g_n(k)\left(\bar{\epsilon}_k-\bar{\epsilon}\right)\right)^2\leqslant 8\max_x g^2(x)\sum_{k=-\infty}^{\infty}|E(\epsilon_0\epsilon_k)|<\infty. \tag{38}$$

(i) For $\delta\geqslant\frac{1}{2}$, since (37) and (38) imply that $II_n+III_n=o_P(1)$, $Q_n(\sigma)$ converges to $N(0,1)$ in distribution.

(ii) For $\delta<\frac{1}{2}$, (36)–(38) and the assumption imply that

$$\begin{aligned}
\sqrt{\frac{4n}{3[nh]}}n^{2\delta-1}Q_n(\sigma)&=\sqrt{\frac{4n}{3[nh]}}n^{2\delta-1}\left(I_n+III_n\right)+\frac{1}{n\sigma^2}\sum_{k=0}^{n-[nh]}g_n^2(k)\\
&\xrightarrow{\mathrm{P}}\frac{1}{\sigma^2}\int_0^1\left(g(x)-\int_0^1 g(x)\,\mathrm{d}x\right)^2\mathrm{d}x.
\end{aligned}$$

This completes the proof. $\square$

Proof of Theorem 3.3. Since (13), (14) and (16) hold, we have

$$
\begin{aligned}
&\left| \max_{0\leqslant k\leqslant n-[nh]} \left\| \Sigma^{-1/2}\left(\hat{\boldsymbol\theta}_k - \hat{\boldsymbol\theta}\right)\right\| - \max_{0\leqslant k\leqslant n-[nh]} \left\| \Sigma^{-1/2}\bar{\boldsymbol\eta}_k \right\| \right| \\
&\quad \leqslant \max_{0\leqslant k\leqslant n-[nh]} \left\| \Sigma^{-1/2}\left(\hat{\boldsymbol\theta}_k - \hat{\boldsymbol\theta} - \bar{\boldsymbol\eta}_k\right)\right\| \\
&\quad \leqslant C_2 \left(\max_{0\leqslant k\leqslant n-[nh]} \|\boldsymbol\Delta_k\| + \left\|\hat{\boldsymbol\theta} - \boldsymbol\theta_0\right\| \right) \\
&\quad = o_P\left(\frac{1}{\sqrt{nh\log h^{-1}}}\right) + O_P\left(\frac{1}{\sqrt{n}}\right)
\end{aligned}
$$

for some positive constant C_2. Hence, the first statement in the theorem follows from Lemma 5.1.

The second statement is similarly proven by Lemma 5.2 and the following fact:

$$
\begin{aligned}
&\left| \max_{0\leqslant k\leqslant n-[nh]} \left\| \Sigma^{-1/2}\left(\hat{\boldsymbol\theta}_k - \hat{\boldsymbol\theta}\right)\right\|_2 - \max_{0\leqslant k\leqslant n-[nh]} \left\| \Sigma^{-1/2}\bar{\boldsymbol\eta}_k \right\|_2 \right| \\
&\quad \leqslant \max_{0\leqslant k\leqslant n-[nh]} \left\| \Sigma^{-1/2}\left(\hat{\boldsymbol\theta}_k - \hat{\boldsymbol\theta} - \bar{\boldsymbol\eta}_k\right)\right\|_2 \\
&\quad \leqslant d \max_{0\leqslant k\leqslant n-[nh]} \left\| \Sigma^{-1/2}\left(\hat{\boldsymbol\theta}_k - \hat{\boldsymbol\theta} - \bar{\boldsymbol\eta}_k\right)\right\|.
\end{aligned}
$$

This completes the proof. □

Proof of Theorem 3.4. By Lemma 5.5, we have

$$
\sqrt{\frac{3[nh]}{4nd}} \left(\sum_{k=0}^{n-[nh]} \left\| \Sigma^{-1/2}\bar{\boldsymbol\eta}_k \right\|_2^2 - \frac{(n-[nh]+1)d}{[nh]} \right) \xrightarrow{\mathrm{d}} N(0,1). \tag{39}
$$

Meanwhile, from (13), (17) and (39), we get

$$
\begin{aligned}
&\left| \sum_{k=0}^{n-[nh]} \left\| \Sigma^{-1/2}\left(\hat{\boldsymbol\theta}_k - \hat{\boldsymbol\theta}\right)\right\|_2^2 - \sum_{k=0}^{n-[nh]} \left\| \Sigma^{-1/2}\bar{\boldsymbol\eta}_k \right\|_2^2 \right| \\
&\quad \leqslant \sum_{k=0}^{n-[nh]} \left\| \Sigma^{-1/2}\left(\boldsymbol\Delta_k - \hat{\boldsymbol\theta} + \boldsymbol\theta_0\right)\right\|_2^2 + 2\left| \sum_{k=0}^{n-[nh]} \bar{\boldsymbol\eta}_k' \Sigma^{-1}\boldsymbol\Delta_k \right| + 2\left| \sum_{k=0}^{n-[nh]} \bar{\boldsymbol\eta}_k' \Sigma^{-1}\left(\hat{\boldsymbol\theta} - \boldsymbol\theta_0\right) \right| \\
&\quad \leqslant C_3^2 \left(\sum_{k=0}^{n-[nh]} \|\boldsymbol\Delta_k\|_2^2 + n\left\|\hat{\boldsymbol\theta} - \boldsymbol\theta_0\right\|_2^2 \right) + C_3 \left(\sum_{k=0}^{n-[nh]} \left\| \Sigma^{-1/2}\bar{\boldsymbol\eta}_k \right\|_2^2 \sum_{k=0}^{n-[nh]} \|\boldsymbol\Delta_k\|_2^2 \right)^{1/2} \\
&\qquad + C_3^2 \left\| \frac{1}{\sqrt{n}} \sum_{k=0}^{n-[nh]} \bar{\boldsymbol\eta}_k \right\|_2 \left\| \sqrt{n}\left(\hat{\boldsymbol\theta} - \boldsymbol\theta_0\right)\right\|_2 \\
&\quad = o_P\left(1/\sqrt{h}\right)
\end{aligned}
$$

for some positive constant C_3, since $n^{-1/2}\sum_{k=0}^{n-[nh]} \bar{\boldsymbol\eta}_k = O_P(1)$. Hence, (39) establishes (18). □

Proof of Theorem 3.5. Since $\{\boldsymbol{\eta}_{1n}, n \geqslant 1\}$ satisfies (15) for some Σ_1 and $0 < \lambda_1 < \frac{1}{2}$ and h satisfies (4) with λ_1,

$$\max_{[nh] \leqslant k \leqslant n-[nh]} \left\| \frac{1}{[nh]} \left(\sum_{i=k+1}^{k+[nh]} \boldsymbol{\eta}_{1i} - \sum_{i=k-[nh]+1}^{k} \boldsymbol{\eta}_{1i} \right) \right\| = O_P\left(\sqrt{\frac{\log h^{-1}}{nh}} \right)$$

by Lemma 5.3. Therefore, we have that from (22)

$$\begin{aligned}
&\max_{[nh] \leqslant k \leqslant [n\tau]-[nh]} \left\| \hat{\boldsymbol{\theta}}_k - \hat{\boldsymbol{\theta}}_{k-[nh]} \right\| \\
&\quad \leqslant \max_{[nh] \leqslant k \leqslant n-[nh]} \left\| \frac{1}{[nh]} \left(\sum_{i=k+1}^{k+[nh]} \boldsymbol{\eta}_{1i} - \sum_{i=k-[nh]+1}^{k} \boldsymbol{\eta}_{1i} \right) \right\| + \max_{0 \leqslant k \leqslant n-[nh]} \|\boldsymbol{\Delta}_{1k}\| \\
&\quad = o_P(1).
\end{aligned}$$

Similarly,

$$\max_{[n\tau]+[nh] \leqslant k \leqslant n-[nh]} \left\| \hat{\boldsymbol{\theta}}_k - \hat{\boldsymbol{\theta}}_{k-[nh]} \right\| = o_P(1),$$

and

$$\hat{\boldsymbol{\theta}}_{[n\tau]} - \hat{\boldsymbol{\theta}}_{[n\tau]-[nh]} = \boldsymbol{\theta}_2 - \boldsymbol{\theta}_1 + o_P(1).$$

Hence,

$$\lim_{n \to \infty} P\left\{ [n\tau] - [nh] \leqslant n\hat{\tau} \leqslant [n\tau] + [nh] \right\} = 1,$$

which entails (23). This completes the proof. □

Acknowledgments

We thank the Associate Editor and the two referees for their valuable suggestions in improving the quality of this paper. This research was supported by Korean Research Foundation Grant 2003-070-C00008.

References

[1] J.M.P. Albin, On extremal theory for stationary processes, Ann. Probab. 18 (1990) 92–128.
[2] P. Bickel, P. Rosenblatt, On some global measures of the deviation of density function estimators, Ann. Stat. 1 (1973) 1071–1095.
[3] R. Bradley, Basic properties of strong mixing conditions, in: E. Eberlein, M.S. Taqqu (Eds.), Dependence in Probability and Statistics, Birkhäuser, Boston, 1986, pp. 303–334.
[4] C.-S.J. Chu, K. Hornick, C.-M. Kuan, The moving-estimates test for parameter stability, Econ. Theory 11 (1995) 699–720.
[5] M. Csörgő, L. Horváth, Limit Theorems in Change-point Analysis, Wiley, New York, 1997.
[6] E. Eberlein, On the strong invariance principles under dependence assumptions, Ann. Probab. 14 (1986) 260–270.
[7] P. Gaenssler, E. Haeusler, On martingale central limit theory, in: E. Eberlein, M.S. Taqqu (Eds.), Dependence in Probability and Statistics, Birkhäuser, Boston, 1986, pp. 303–334.
[8] C. Inclán, G.C. Tiao, Use of cumulative sums of squares for retrospective detection of changes of variances, J. Am. Stat. Assoc. 89 (1994) 913–923.
[9] V.K. Jandhyala, I.B. MacNeill, Tests for parameter changes at unknown times in linear regression models, J. Stat. Plann. Infererence 27 (1991) 291–316.

[10] C.M. Kuan, K. Hornik, The generalized fluctuation test: a unifying view, Econ. Rev. 14 (1995) 135–161.
[11] J. Kuelbs, W. Philipp, Almost sure invariance principles for partial sums of mixing B-valued random variables, Ann. Probab. 8 (1980) 1003–1036.
[12] S. Lee, J. Ha, O. Na, S. Na, The cusum test for parameter change in time series models, Scand. J. Stat. 30 (2003) 781–796.
[13] S. Lee, O. Na, Test for parameter change in stochastic processes based on conditional least-squares estimator, J. Multivariate Anal. 93 (2005) 375–393.
[14] S. Lee, O. Na, S. Na, On the cusum of squares test for variance change in nonstationary and nonparametric time series models, Ann. Inst. Statist. Math. 55 (2003) 467–485.
[15] P. Révész, Random Walk in Random and Non-random Environments, World Scientific, Singapore, 1990.
[16] S.M. Tang, I.B. MacNeill, The effect of serial correlation on tests for parameter change at unknown time, Ann. Stat. 21 (1993) 552–575.
[17] C.S. Withers, Conditions for linear processes to be strong-mixing, Z. Wahrscheinlichkeit. Verw. Geb. 57 (1981) 477–480.
[18] D. Picard, Testing and estimating change-points in time series, Adv. Appl. Probab. 17 (1985) 841–867.